INEEL/CON-02-00161
PREPRINT

Idaho National Engineering and Environmental Laboratory

Machine Vision for Object Detection and Profiling in an Unstructured Environment

R. Kinoshita
M. Walton

August 4, 2002 – August 8, 2002

Spectrum 2002

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint should not be cited or reproduced without permission of the author.

This document was prepared as a account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights. The views expressed in this paper are not necessarily those of the U.S. Government or the sponsoring agency.

Machine Vision for Object Detection and Profiling in an Unstructured Environment

Robert Kinoshita
Idaho National Engineering and Environmental Laboratory
2525 N. Fremont Av.
MS 2220
Idaho Falls, ID 83415-2220
(208) 526-4768
bobk@inel.gov

Miles Walton
Idaho National Engineering and Environmental Laboratory
2525 N. Fremont Av.
MS 2220
Idaho Falls, ID 83415-2220
(208) 526-3087
mwalton@inel.gov

Abstract

The Handling and Sorting System for 55-Gallon Drums (HANDSS-55) is a DOE project to develop an automated method for retrieving items that are not acceptable at the Waste Isolation Pilot Plant (WIPP) from 55-gallon drums of low-level waste. The HANDSS-55 is a modular system that opens drums, sorts the waste, and then repackages the remaining waste in WIPP compliant barrels. The Sorting Station module relies on a non-contact measurement system, the StarCam? [a], to quickly provide a 3D profile of the sorting area. It then analyses the 3D profile and a color image to determine the position and orientation of an operator selected waste item. The item is then removed from the sorting area by a robotic arm. The use of both image and profile information for object determination provides a fast, effective method of finding and retrieving selected objects in the unstructured environment of the sorting module.

I. INTRODUCTION

At many locations around the United States waste items contaminated with transuranic elements have been stored for decades. The Waste Isolation Pilot Plant (WIPP) is now available to permanently store these materials. There are many requirements imposed on waste sent to WIPP. At the Savannah River Site (SRS), thousands of 55-gallon drums of such waste await sorting and inspection to ensure that all materials conform to the acceptance criteria. Currently, this sorting is done by hand at a rate of one drum per day. The hand sorting entails at least five people in the process, three of which are suited up in anti-contamination clothing, working through ports in a glove box. The Idaho National Engineering and Environmental Laboratory (INEEL) is developing an automated alternative called the Handling and Sorting System for 55-Gallon Drums (HANDSS-55). This system provides a high degree of automation in processing 55-gallon drums of waste for shipment and storage at the WIPP. Waste drums with non-compliant items, identified through radiography, are automatically opened, the contents are placed on a sorting table and any non-compliant items are removed. The automatic removal of these items from the table is accomplished by the use of a stereo vision measurement system and sophisticated image analysis software. This provides the object size and position on the table. The object is then removed from the table by a robotic arm. A key to the success of the HANDSS-55 is making the process of picking up items as fast and accurate as possible. This paper discusses the stereo vision measurement system and the image analysis software used for object detection.

II. WORKING DESCRIPTION

The HANDSS-55 consists of several modules used to open, sort, and repackage low level waste drums in preparation for shipment to WIPP. The Sorting Station module provides a high level of automation that allows an operator to remove non-compliant items from the waste stream. The equipment for the module consists of a sorting table, a robotic arm with various end effectors, a stereo vision measurement system, and electronic and pneumatic sensors and control devices. These devices are integrated by a network of computers running Microsoft

[a] PRODUCT DISCLAIMER
References herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government, any agency thereof, or any company affiliated with the Idaho National Engineering and Environmental Laboratory.

Windows 2000 that provide an operator interface and distributed equipment control in a client/server setup. Drum opening and waste repackaging are conducted in other modules of the HANDSS-55. This paper will discuss in detail the operator control interface client software, the stereo vision measurement system with its software server, and the image analysis software server.

The Sorting Station Operator Control Client is a software client that provides the operator with a live video feed of the sorting table and control buttons for system actions. The operator can interact with the program through a touch screen monitor or with a keyboard and mouse. The live video feed comes from a stereo vision measurement system, the StarCam? , mounted approximately three meters above the sorting table. To select an object for removal from the table, the operator touches the object on the screen image and identifies a region of interest (ROI) completely surrounding the object. The StarCam? unit then performs a measurement scan and provides the control client with an array of spatial coordinates for points within the selected region. These points, along with the image, is passed to the image analysis server which returns to the control program an outline and position data for the object. The StarCam? measurement system and the image analysis will be described in more detail later. Once the coordinates of the object have been determined, the operator selects an appropriate end effector and an appropriate receptacle for the object, which is then removed from the table. This process is repeated until all non-compliant items are removed from the table. At that point the table is automatically cleared and remaining waste is repackaged. The Sorting Station Operator Control Client was developed using Microsoft Visual Basic and uses the industry standard Component Object Model/ Distributed Component Object Model (COM/DCOM) interface definition to interact with the numerous control servers on the system.

III. STEREO VISION MEASUREMENT SYSTEM

The heart of the system is the StarCam? FW 3D Precision Measurement System, shown in Figure 1. It is a unique, non-contact distance measurement device developed by VX Technologies, Inc. The StarCam? uses stereo cameras and structured lighting, consisting of arrays of white dots projected by an LCD projector, to calculate distance. The StarCam? is controlled by a computer system that communicates with the StarCam? over an IEEE-1394 bus. This bus, commonly called “Fire Wire”, is a high-speed communication standard that provides digital transmission of image and control data. The StarCam? is controlled by a COM compliant server that accepts commands from the control client and provides a video stream and ranging data.

Figure 1 - The StarCam Precision Measurement System. Courtesy VX Technologies Inc.

The StarCam? performs sterographic rangefinding using a pair of CCD digital video cameras. Stereo rangefinding is simple in concept but difficult in practice. By finding the same target point in each image the distance to the point can be calculated using triangulation. The StarCam? uses the left camera for the reference image. Once a point is identified in both the left and right images the distance is calculated between parallel planes perpendicular to the optic axis of the left camera, one plane passing through the point and the other plane at the front plate of the StarCam? . After the distance is calculated the position of the point is returned as two coordinate sets. The distance coordinate, (X, Y, Z) has its origin (0,0,0) occurring at the point where the optic axis of the left camera passes through the front plate of the StarCam? . The image coordinate, (X, Y) is the pixel position of the point on the left camera image and has its origin (0,0) at the upper left of the image.

The ability of the StarCam? software to find target points for distance calculations relies on being able to correlate image features in the two images. The correlation relies on detecting regions of contrast that match. One consequence of this is that a scene with minimal contrast cannot be measured. In order to provide a reliable degree of contrast, structured lighting, consisting of patterns of white dots to provide a good edge, are projected on the surface to be measured. The pattern currently used provides a 2.5-cm spacing between dots.

With stereoscopic rangefinding a measurement is always made to a region and not a point. The minimum region that can be measured has a width of about 10 pixels. If a feature is found in a region in the reference image then a search is made in the secondary image for a corresponding feature. If the feature is found then the distance calculation can be made. Exposure in the images is also a factor in achieving good contrast for image features. The StarCam? software automatically steps through a range of exposure levels seeking the level that provides the best contrast in the region.

A target point can only be measured if it is found in both images. This depends on the alignment of the cameras and the distance of the point from the StarCam? . The alignment of the cameras is calibrated when the unit is manufactured and is not field adjustable. However, the user of the StarCam? can specify the range for the distances that the unit will attempt to measure. The broader the range to be measured the longer the search for matching regions will take. The Sorting Station uses a range of 2.5 to 3 meters. If an item of interest is outside this range then it will not reliably be measured. A region of interest (ROI) in the image may also be specified to decrease computation time. When an ROI has been specified the StarCam? software searches for matching features in only that portion of the image.

The StarCam? continuously self calibrates through the use of an internal optical reference. The optical reference provides a pair of artificial target beams that have a stable parallax angle. Essentially, there exists a target point at a precise distance in front of the StarCam? available as a reference. The reference point is used by the software to correct for false image separations that may be caused by thermal changes, shocks to the unit, or changes in focus in the cameras. The StarCam? is able to resolve target images separations to the level of $1/100^{th}$ the pixel width.

IV. IMAGE ANALYSIS SYSTEM

A. Problem Description.

The imaging system is very important to waste sorting operations because of the need to quickly and accurately move objects of indeterminate size, shape, and weight. The sorting station XYZ Deployment System (DS) has been designed to be able to reach any location in the entire work area (approximately 1 x 3 x 1meters) within five seconds. A variety of end effectors can be used to lift objects, up to a maximum of over 200 kg. In addition, the workspace is cluttered—especially just after a new drum of waste to be processed has been dumped onto the table. Figure 2 shows how such a workspace appears from overhead to the StarCam? .

Figure 2 - Sorting Table with Several Objects

In contrast to today's typical image processing task, the software described in this paper is used to look for unknown objects in an unknown environment. Today's available commercial software is designed around looking for predetermined objects as part of a manufacturing, sorting, or counting operation. Current applications include rapid quality inspection of large numbers of manufactured parts and checking to make sure that components are properly placed on a circuit board. Instead of finding out what an object is, we instead must separate said item from the clutter around it, and provide sufficient information to successfully pick it up. Since the deployment system has a lot of power, equipment damage is of particular concern. Therefore it is important to have as much information about the location and geometry of a target object as possible. This is particularly true of the height: Overshooting the top surface of an object during a pick-up operation can damage the object and/or end-effector. Figure 3 shows a soda can before and after the Z-Mast was used to apply about 3500 Newtons of force from above—which simulates such an event. The image analysis module is designed to provide the best possible estimate of object location, distance, and boundary information for arbitrary operator-selected regions on the sorting table.

Figure 3 – Object Size Reduction is Not the Goal!

B. System Inputs

Computationally locating an arbitrary item is a difficult problem to solve. Even an operator using the video image to select objects can make a mistake in identifying the boundary of an object. Therefore, every available piece of information must be considered together to make the best possible object pickup determinations. To do this, the sorting station makes use of the HANDSS-55 system image processing software, which is implemented as a Distributed Component Object Model (DCOM) module. Information described in this section is collected by the control program and made available to the module when it is asked to locate an object for pickup.

1.Object Touch Point and Bounding Rectangle

During the automated waste sorting process, objects are selected for pickup individually. The sorting station operator doing this uses the live video feed from the StarCam™ supplemented by the auxiliary glovebox cameras to determine whether or not an object needs to be moved. The touch screen or mouse is then used to draw a rectangle around the object of interest. This is done by touching the center of the object as shown on the StarCam™ video and drawing a corner outward, forming a "rubber band" box that encloses the target. This Region of Interest (ROI) is used to narrow the area for which distance information is collected, which saves processing time.

2. Video Information

Each time a search is initiated, one color still video frame of the entire sorting table is provided to the image analysis module. This information is provided in the form of three sets of 8-bit data (per pixel)—one each for red, green, and blue colors. These are combined in the image processing module to form a color picture. Figure 1 (above) is one of these pictures written as a file. The arrangement of color dots in this image can (and will!) be analyzed to discern patterns that may or may not indicate the boundaries of an object.

3. StarCam™ Range Information

Range information is provided as an array of point, each of which contains the following five components: X- Y- and Z- displacement of a location in space from the camera face in physical units (meters in this case). The X- and Y- coordinates are also provided in pixel coordinates, which allows these points to be correlated to the video image. These points taken together can be represented as a shaded model to express differences in height across the sorting table. Figure 4 is such a representation of the table arrangement already presented in Figure 2, shown at an oblique angle from vertical to make the differences in height easier to see. Such views throughout this paper are constructed using the MetaFlash™ viewer program provided by the Viewpoint Corporation.

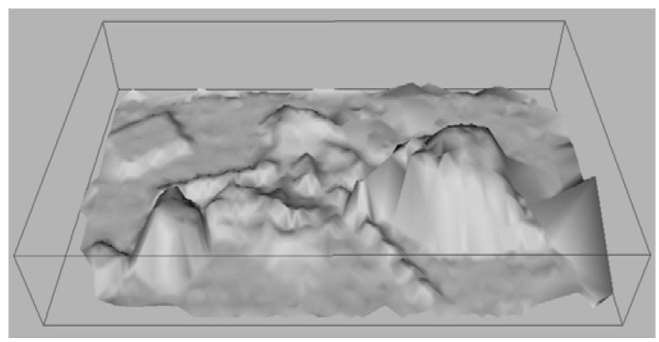

Figure 4 – MetaFlash™ Visual Representation of Range Data

One well-known weakness of using standard video cameras is lack of depth perception. The piece of wood shown on the sorting table in Figure 5 illustrates this. From above, it appears that the wood is lying flat, yet the StarCam™ data reveals that the board is actually inclined.

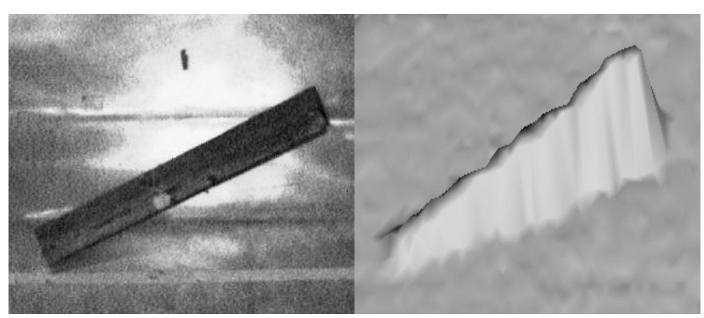

Figure 5 - This board looks flat...or is it?

The range data is supplied as an array of points in no particular order, and with no guarantee that particular dots are included. For example, surfaces of certain colors and textures do not yield StarCam™ points as well as others. Therefore we must find some way to estimate the range between points. The image processing software and the MetaFlash™ viewer use the same mesh-based method to do this, which is discussed below.

C. Outputs

During the analysis process, many useful pieces of information are generated and made available to the Sorting Station control program for use in positioning the Z-Mast and its attached gripper effectively.

1. Mesh Estimate

To provide a way to estimate a distance anywhere in the ROI, all available StarCam™ data points are used to generate a mesh of triangles. This technique, Voroni Tesselation, is very useful for rendering an image so that it has a three-dimensional look (i.e. MetaFlash viewer, thousands of computer games). A surface is then fitted to points interpolated on these triangles and formed into an image that represents varying heights by pixel value (brighter is higher). Figure 6 shows the image of a metal beam and its corresponding mesh.

Figure 6 - Video Image and Mesh Example (Brighter is Higher)

A way is provided for the sorting station program to inquire of the height estimate derived from this image at any point of the ROI. An estimate of the quality of the result is also available, which is dependent on the number and distance to nearby StarCam™ data points used to generate the mesh. Using Figure 6 as an example, the beam on the left appears to be continuous yet the mesh on the right shows a “hole” in the middle of it. This was because no distance data point was found on the beam in that area so nearby low points on the table surface drew the mesh lower.

2. Calculated Object Geometry

Once the object finder has finished its search, several parameters indicating the shape and orientation are returned to the client. To provide a visual indication of the result, an array of pixels coordinates are returned that form an outline when overlaid on the object selection display. The centroid and the endpoints of the major and minor axes of this outline are also computed. The ratio of the lengths of these lines can indicate general object shape (i.e. Long and thin vs. square). This information, combined with distance measurements, are essential for making good decisions about how to choose a gripper and position it to pick up an object.

3. Trace Image

For developmental purposes, the image analysis routine can be directed to create and write a trace image into a standard Windows bitmap (.BMP) file for later examination. Figure 7 is an example of such an image. When an analysis cycle starts the trace is started using a copy of the sorting table raw color video. As the process continues additional information is drawn into this image as needed in various steps to guide development and tuning. This capability was very helpful in identifying faults during the initial development of this software. When not needed, the generation of this file can be disabled to decrease processing time, which is on the order of one second per cycle.

Figure 7 - Object Finder Trace Image Example

D. Analysis Methodology

1. Image Analysis Library

Image processing is a rapidly growing field that has a great amount of research interest and commercial product offerings to assist in development. Most of the commercial products are structured around the rapid inspection of like objects and the recognition of characters and bar codes in images. There are also several libraries providing low-level functions that are for incorporation into software like that of this project. The Open Source Computer Vision Library (OpenCV) for Intel® architectures is one such library.

OpenCV was selected for use in this project based on several factors. First, it is backed by the Intel Corporation and targeted directly at the platform already selected for this project (Microsoft Windows™ 2000 operating system running on Intel processors). The basic routines in this library are optimized for performance when so used. This library is also undergoing current development and use in both university and commercial settings as demonstrated by an active user community including frequent participation by the library developers. The price is excellent—free for use within the license agreement, and the source code is provided with the library. Having this code has been useful several times to track down problems during system development. Image processing algorithms in many research areas have been implemented in OpenCV. Feature detection, shape analysis, and object segmentation/recognition are functional areas that have proved most useful for analysis. Basic image functions such as display, file input/output, and color conversion have also been invaluable.

2. Assumptions made during analysis

? The center of the ROI is located on what the operator is trying to select.

? The object occupies a significant part of the ROI.

? The object will be (at least mostly) the highest object in the ROI.

? Operators will grow familiar with how the image processing algorithm uses their selection and adjust it to improve performance.

? Objects may be arranged in any orientation, and will very often be tilted.

? Objects may be partially covered by other objects.

3. Contour Extraction and Matching

A *contour* is an image processing concept that provides a way to describe a set of points that circumscribe an object of interest. OpenCV provides data structures and functions to search for, extract, analyze and draw these. The video image input and StarCam™ data are at first examined separately for contours. In the video case, the image is converted to grey scale (non-color) format and processed with a routine that highlights strong changes in image intensity moving away from the center of the ROI. The result is then selectively thresholded and searched for the dominant contour. This family of curves is stored in memory after removing obviously invalid results (generally excessively large or small contours). A similar family of curves is generated for the StarCam™ distance data looking for sudden high-to-low transitions in the height. The low-to-high transitions are ignored for this purpose because of the special case of objects that have holes in them. Figure 8 is an example of such an object (a roll of duct tape). Once the two families of contours are ready, they are compared against each other using several comparison methods available in the OpenCV library. The pair of contours (one each from video and StarCam™ data) that scores highest is then passed on to the geometry analysis routine for additional decision making based on contour geometry.

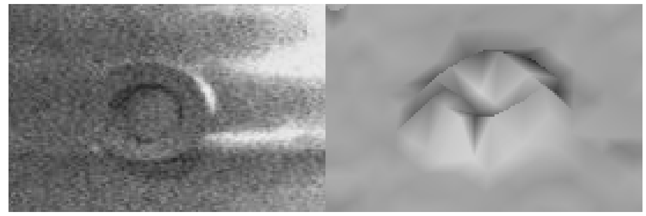

Figure 8 - Example Object with a Hole in it

4. Contour Geometry Analysis

The two contours representing the best match are then analyzed to determine their geometry. OpenCV contour-processing functions are used to calculate the area and area-based centroid of each outline. The results and their contours are also shown on the trace image (example Figure 7, above) for later evaluation as needed. These values are then compared to the ROI size and center to decide which of the two that will be returned as a result. This is done (in keeping with the above assumptions) to bias system decisions against extremely small or large contours, and those that are centered far away from the touch point.

5. End-Effector Finger Positioning

The more suitable of the two contours (whether video- or distance-based) is then further processed to provide specific information needed to position the XYZ DS and end effector to pick up the object. Major and minor axes of the shape are provided to indicate general object orientation. Finger positions to prepare for a grasp are suggested in case the multi-purpose end effector is selected for use. These values have also been indicated on the trace image. Most important of these is StarCam™ distance information, which is given both for the centroid (pick-up point) and for the surrounding surface. This is needed to position for a successful grab, and to prevent positioning so low that something gets crushed (remembering the pop can in Figure 3).

V. RESULTS

The success rate for object detection is around 98% for discrete objects on the table and 80% for objects on a cluttered table. The distance measurement and image analysis takes approximately 15 seconds to complete, depending on the size of the ROI. Research into improving the image analysis process is being conducted using simulated waste objects to provide increasingly difficult object discrimination. As a backup to automatic analysis, software controls are given to the operator to allow the repositioning of the outline and pickup points manually on the screen prior to actually picking up the object.

VI. CONCLUSION

The use of the StarCam? stereo vision measurement system and sophisticated image analysis tools allows for the rapid determination of object position in an unstructured environment. A simple user interface and advanced automation allow the rapid sorting and removal of non-compliant items from barrels of low level waste bound for WIPP for permanent disposal. The automation provided by HANDSS-55 will increase waste throughput and decrease personnel exposure in waste sorting operations at Savannah River Site.

VII. ACKNOWLEDGEMENTS

Work supported by the TRU and Mixed Waste Focus Area of the U.S. Department of Energy, Office of Science and Technology EM 50, under DOE contract number DE-AC07-99ID13727.

VIII. REFERENCES

1. VX Optronics Corp., *StarCam? Module X^2Z Users Manual,* p.15, VX Optronics Corp., Calgary, Alberta, Canada, (1999)
2. Intel Corporation, *Open Source Computer Vision Library, Reference Manual*, Intel Corporation, http://www.intel.com/research/mrl/research/opencv/OpenCVReferenceManual.pdf, (2001)